

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
8 July 2004 (08.07.2004)

PCT

(10) International Publication Number  
**WO 2004/056967 A2**

- (51) International Patent Classification<sup>7</sup>: **C12N**
- (21) International Application Number:  
PCT/US2003/040987
- (22) International Filing Date:  
18 December 2003 (18.12.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/434,601 18 December 2002 (18.12.2002) US
- (71) Applicant (for all designated States except US): **AGRI-NOMICS LLC** [US/US]; 16160 S.W. Upper Boones Ferry Road, Portland, OR 97224-7744 (US).
- (72) Inventors; and
- (73) Inventors/Applicants (for US only): **LIGHTNER, Jonathan** [US/US]; 28750 S. Heisinger Lane, Mulino, OR 97042 (US). **COATE, Jeremy, E.** [US/US]; 3202 S.E. 10th Avenue, Portland, OR 97202 (US). **CLENDEN-NEN, Stephanie, K.** [US/US]; P.O. Box 6332, Kingsport, TN 37663 (US). **FEDERSPIEL, Nancy, Anne** [US/US]; 1345 Altschul Avenue, Menlo Park, CA 94025 (US).
- (74) Agents: **BRUNELLE, Jan, P. et al.; EXELIXIS, INC.**, 170 Harbor Way, P.O. Box 511, South San Francisco, CA 94083-0511 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: GENERATION OF PLANTS WITH ALTERED OIL CONTENT

(57) Abstract: The present invention is directed to plants that display an altered oil content phenotype due to altered expression of an aconitase nucleic acid. The invention is further directed to methods of generating plants with an altered oil content phenotype.

## GENERATION OF PLANTS WITH ALTERED OIL CONTENT

### REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. provisional patent application no.

5 60/434,601 filed 12/18/2002, the contents of which are hereby incorporated by reference

### BACKGROUND OF THE INVENTION

The ability to manipulate the composition of crop seeds, particularly the content and composition of seed oils, has important applications in the agricultural industries, relating both to processed food oils and to oils for animal feeding. Seeds of agricultural crops contain a variety of valuable constituents, including oil, protein and starch.

Industrial processing can separate some or all of these constituents for individual sale in specific applications. For instance, nearly 60% of the US soybean crop is crushed by the soy processing industry. Soy processing yields purified oil, which is sold at high value, while the remainder is sold principally for lower value livestock feed (US Soybean Board, 2001 Soy Stats). Canola seed is crushed to produce oil and the co-product canola meal (Canola Council of Canada). Nearly 20% of the 1999/2000 US corn crop was industrially refined, primarily for production of starch, ethanol and oil (Corn Refiners Association).

Thus, it is often desirable to maximize oil content of seeds. For instance, for processed oilseeds such as soy and canola, increasing the absolute oil content of the seed will increase the value of such grains. For processed corn it may be desired to either increase or decrease oil content, depending on utilization of other major constituents. Decreasing oil may improve the quality of isolated starch by reducing undesired flavors associated with oil oxidation. Alternatively, in ethanol production, where flavor is unimportant, increasing oil content may increase overall value. In many fed grains, such as corn and wheat, it is desirable to increase seed oil content, because oil has higher energy content than other seed constituents such as carbohydrate. Oilseed processing, like most grain processing businesses, is a capital-intensive business; thus small shifts in the distribution of products from the low valued components to the high value oil component can have substantial economic impacts for grain processors.

Biotechnological manipulation of oils can provide compositional alteration and improvement of oil yield. Compositional alterations include high oleic soybean and corn oil (US Pat Nos 6,229,033 and 6,248,939), and laurate-containing seeds (US Pat No 5,639,790), among others. Work in compositional alteration has predominantly focused

on processed oilseeds but has been readily extendable to non-oilseed crops, including corn. While there is considerable interest in increasing oil content, the only currently practiced biotechnology in this area is High-Oil Corn (HOC) technology (DuPont, US Pat. No. 5,704,160). HOC employs high oil pollinators developed by classical selection  
5 breeding along with elite (male-sterile) hybrid females in a production system referred to as TopCross. The TopCross High Oil system raises harvested grain oil content in maize from ~3.5% to ~7%, improving the energy content of the grain.

While it has been fruitful, the HOC production system has inherent limitations. First, the system of having a low percentage of pollinators responsible for an entire field's  
10 seed set contains inherent risks, particularly in drought years. Second, oil contents in current HOC fields have plateaued at about 9% oil. Finally, high-oil corn is not primarily a biochemical change, but rather an anatomical mutant (increased embryo size) that has the indirect result of increasing oil content. For these reasons, an alternative high oil strategy, particularly one that derives from an altered biochemical output, would be  
15 especially valuable.

The most obvious target crops for the processed oil market are soy and rapeseed, and a large body of commercial work (e.g., US Pat No: 5,952,544; PCT application WO9411516) demonstrates that *Arabidopsis* is an excellent model for oil metabolism in these crops. Biochemical screens of seed oil composition have identified *Arabidopsis*  
20 genes for many critical biosynthetic enzymes and have led to identification of agronomically important gene orthologs. For instance, screens using chemically mutagenized populations have identified lipid mutants whose seeds display altered fatty acid composition (Lernieux B, et al. 1990, Theor Appl Genet 80, 234-240; James DW and Dooner HK (1990) Theor Appl Genet 80, 241-245). T-DNA mutagenesis screens  
25 (Feldmann *et al.*, *Science* 243: 1351-1354, 1989) that detected altered fatty acid composition identified the omega 3 desaturase (*FAD3*) and delta-12 desaturase (*FAD2*) genes (US Pat No 5952544; Yadav NS et al. (1993) Plant Physiol 103, 467-476; Okuley et al., Plant Cell. 1994 Jan;6(1):147-58). A screen which focused on oil content rather than oil quality, analyzed chemically-induced mutants for wrinkled seeds or altered seed  
30 density, from which altered seed oil content was inferred (Focks N and Benning C, Plant Physiol 118:91-101, 1998). Another screen, designed to identify enzymes involved in production of very long chain fatty acids, identified a mutation in the gene encoding a diacylglycerol acyltransferase (DGAT) as being responsible for reduced triacyl glycerol accumulation in seeds (Katavic V et al, Plant Physiol. 1995 May;108(1):399-409). It was

further shown that seed-specific over-expression of the DGAT cDNA was associated with increased seed oil content (Jako et al., Plant Physiol. 2001 Jun;126(2):861-74).

### SUMMARY OF THE INVENTION

5       The invention provides a transgenic plant having a high oil phenotype. The transgenic plant comprises a transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that encodes an aconitase polypeptide. In preferred embodiments, the transgenic plant is selected from the group consisting of rapeseed, soy, corn, sunflower, cotton, cocoa, safflower, oil palm, coconut palm, flax,  
10    castor and peanut. The invention further provides a method of producing oil comprising growing the transgenic plant and recovering oil from said plant.

      The transgenic plant of the invention is produced by a method that comprises introducing into progenitor cells of the plant a plant transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that encodes an  
15    aconitase polypeptide, and growing the transformed progenitor cells to produce a transgenic plant, wherein the aconitase polynucleotide sequence is expressed causing the high oil phenotype.

### DETAILED DESCRIPTION OF THE INVENTION

#### 20    Definitions

      Unless otherwise indicated, all technical and scientific terms used herein have the same meaning as they would to one skilled in the art of the present invention.

      Practitioners are particularly directed to Sambrook *et al.* Molecular Cloning: A Laboratory Manual (Second Edition), Cold Spring Harbor Press, Plainview, N.Y., 1989, and Ausubel  
25    FM *et al.* Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., 1993, for definitions and terms of the art. It is to be understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary.

      As used herein, the term "vector" refers to a nucleic acid construct designed for  
30    transfer between different host cells. An "expression vector" refers to a vector that has the ability to incorporate and express heterologous DNA fragments in a foreign cell. Many prokaryotic and eukaryotic expression vectors are commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

A "heterologous" nucleic acid construct or sequence has a portion of the sequence that is not native to the plant cell in which it is expressed. Heterologous, with respect to a control sequence refers to a control sequence (*i.e.* promoter or enhancer) that does not function in nature to regulate the same gene the expression of which it is currently  
5 regulating. Generally, heterologous nucleic acid sequences are not endogenous to the cell or part of the genome in which they are present, and have been added to the cell, by infection, transfection, microinjection, electroporation, or the like. A "heterologous" nucleic acid construct may contain a control sequence/DNA coding sequence combination that is the same as, or different from a control sequence/DNA coding sequence  
10 combination found in the native plant.

As used herein, the term "gene" means the segment of DNA involved in producing a polypeptide chain, which may or may not include regions preceding and following the coding region, *e.g.* 5' untranslated (5' UTR) or "leader" sequences and 3' UTR or "trailer" sequences, as well as intervening sequences (introns) between individual coding segments  
15 (exons) and non-transcribed regulatory sequence.

As used herein, "recombinant" includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid sequence or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found in identical form within the native (non-recombinant) form of the cell or  
20 express native genes that are otherwise abnormally expressed, under expressed or not expressed at all as a result of deliberate human intervention.

As used herein, the term "gene expression" refers to the process by which a polypeptide is produced based on the nucleic acid sequence of a gene. The process includes both transcription and translation; accordingly, "expression" may refer to either a  
25 polynucleotide or polypeptide sequence, or both. Sometimes, expression of a polynucleotide sequence will not lead to protein translation. "Over-expression" refers to increased expression of a polynucleotide and/or polypeptide sequence relative to its expression in a wild-type (or other reference [*e.g.*, non-transgenic]) plant and may relate to a naturally-occurring or non-naturally occurring sequence. "Ectopic expression" refers to  
30 expression at a time, place, and/or increased level that does not naturally occur in the non-altered or wild-type plant. "Under-expression" refers to decreased expression of a polynucleotide and/or polypeptide sequence, generally of an endogenous gene, relative to its expression in a wild-type plant. The terms "mis-expression" and "altered expression" encompass over-expression, under-expression, and ectopic expression.

The term "introduced" in the context of inserting a nucleic acid sequence into a cell, means "transfection", or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid sequence into a eukaryotic or prokaryotic cell where the nucleic acid sequence may be incorporated into the genome of the cell (for example, chromosome, plasmid, plastid, or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (for example, transfected mRNA).

As used herein, a "plant cell" refers to any cell derived from a plant, including cells from undifferentiated tissue (*e.g.*, callus) as well as plant seeds, pollen, progagules and embryos.

As used herein, the terms "native" and "wild-type" relative to a given plant trait or phenotype refers to the form in which that trait or phenotype is found in the same variety of plant in nature.

As used herein, the term "modified" regarding a plant trait, refers to a change in the phenotype of a transgenic plant relative to the similar non-transgenic plant. An "altered oil content phenotype" refers to measurable phenotype of a genetically modified plant, where the plant displays a statistically significant increase or decrease in overall oil content (*i.e.*, the percentage of seed mass that is oil), as compared to the similar, but non-modified plant. A high oil phenotype refers to an increase in overall oil content.

As used herein, a "mutant" polynucleotide sequence or gene differs from the corresponding wild type polynucleotide sequence or gene either in terms of sequence or expression, where the difference contributes to a modified plant phenotype or trait. Relative to a plant or plant line, the term "mutant" refers to a plant or plant line which has a modified plant phenotype or trait, where the modified phenotype or trait is associated with the modified expression of a wild type polynucleotide sequence or gene.

As used herein, the term "T1" refers to the generation of plants from the seed of T0 plants. The T1 generation is the first set of transformed plants that can be selected by application of a selection agent, *e.g.*, an antibiotic or herbicide, for which the transgenic plant contains the corresponding resistance gene. The term "T2" refers to the generation of plants by self-fertilization of the flowers of T1 plants, previously selected as being transgenic. T3 plants are generated from T2 plants, etc. As used herein, the "direct progeny" of a given plant derives from the seed (or, sometimes, other tissue) of that plant and is in the immediately subsequent generation; for instance, for a given lineage, a T2 plant is the direct progeny of a T1 plant. The "indirect progeny" of a given plant derives from the seed (or other tissue) of the direct progeny of that plant, or from the seed (or

other tissue) of subsequent generations in that lineage; for instance, a T3 plant is the indirect progeny of a T1 plant.

As used herein, the term "plant part" includes any plant organ or tissue, including, without limitation, seeds, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores. Plant cells can be obtained from any plant organ or tissue and cultures prepared therefrom. The class of plants which can be used in the methods of the present invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledenous and dicotyledenous plants.

As used herein, "transgenic plant" includes a plant that comprises within its genome a heterologous polynucleotide. The heterologous polynucleotide can be either stably integrated into the genome, or can be extra-chromosomal. Preferably, the polynucleotide of the present invention is stably integrated into the genome such that the polynucleotide is passed on to successive generations. A plant cell, tissue, organ, or plant into which the heterologous polynucleotides have been introduced is considered "transformed", "transfected", or "transgenic". Direct and indirect progeny of transformed plants or plant cells that also contain the heterologous polynucleotide are also considered transgenic.

#### **Identification of Plants with an Altered Oil Content Phenotype**

Transgenic plants were produced to over-express various genes encoding enzymes of the glyoxylate pathway and seed from the T1 transgenic plants were tested for a high oil phenotype. It was discovered that overexpression of aconitase (At4g35830; GI# 13124706) confers an altered oil content phenotype (specifically, a high seed oil phenotype). Accordingly, aconitase genes and/or polypeptides may be employed in the development of genetically modified plants having a modified oil content phenotype. Aconitase genes may be used in the generation of oilseed crops that provide improved oil yield from oilseed processing and in the generation of feed grain crops that provide increased energy for animal feeding. Aconitase genes may further be used to increase the oil content of specialty oil crops, in order to augment yield of desired unusual fatty acids. Transgenic plants that have been genetically modified to express aconitase can be used in the production of oil, wherein the transgenic plants are grown, and oil is obtained from plant parts (e.g. seed) using standard methods.

### Aconitase Nucleic Acids and Polypeptides

*Arabidopsis* aconitase nucleic acid is provided in SEQ ID NO:1 and in Genbank entry GI#30690504. The corresponding protein sequence is provided in SEQ ID NO:2 and in GI#13124706. Nucleic acids and/or proteins that are orthologs or paralogs of *Arabidopsis* aconitase, are described in Example 2 below.

As used herein, the term "aconitase polypeptide" refers to a full-length aconitase protein or a fragment, derivative (variant), or ortholog thereof that is "functionally active," meaning that the protein fragment, derivative, or ortholog exhibits one or more of the functional activities associated with the polypeptide of SEQ ID NO:2. In one preferred embodiment, a functionally active aconitase polypeptide causes an altered oil content phenotype when mis-expressed in a plant. In a further preferred embodiment, mis-expression of the aconitase polypeptide causes a high oil phenotype in a plant. In another embodiment, a functionally active aconitase polypeptide is capable of rescuing defective (including deficient) endogenous aconitase activity when expressed in a plant or in plant cells; the rescuing polypeptide may be from the same or from a different species as that with defective activity. In another embodiment, a functionally active fragment of a full length aconitase polypeptide (i.e., a native polypeptide having the sequence of SEQ ID NO:2 or a naturally occurring ortholog thereof) retains one or more of the biological properties associated with the full-length aconitase polypeptide, such as catalytic activity. An aconitase fragment preferably comprises an aconitase domain, such as a C- or N-terminal or catalytic domain, among others, and preferably comprises at least 10, preferably at least 20, more preferably at least 25, and most preferably at least 50 contiguous amino acids of an aconitase protein. Functional domains can be identified using the PFAM program (Bateman A et al., 1999 Nucleic Acids Res 27:260-262; website at pfam.wustl.edu). Functionally active variants of full-length aconitase polypeptides or fragments thereof include polypeptides with amino acid insertions, deletions, or substitutions that retain one or more of the biological properties associated with the full-length aconitase polypeptide. In some cases, variants are generated that change the post-translational processing of an aconitase polypeptide. For instance, variants may have altered protein transport or protein localization characteristics or altered protein half-life compared to the native polypeptide.

As used herein, the term "aconitase nucleic acid" encompasses nucleic acids with the sequence provided in or complementary to the sequence provided in SEQ ID NO:1, as well as functionally active fragments, derivatives, or orthologs thereof. An aconitase



nucleic acid of this invention may be DNA, derived from genomic DNA or cDNA, or RNA.

In one embodiment, a functionally active aconitase nucleic acid encodes or is complementary to a nucleic acid that encodes a functionally active aconitase polypeptide.

5 Included within this definition is genomic DNA that serves as a template for a primary RNA transcript (i.e., an mRNA precursor) that requires processing, such as splicing, before encoding the functionally active aconitase polypeptide. An aconitase nucleic acid can include other non-coding sequences, which may or may not be transcribed; such sequences include 5' and 3' UTRs, polyadenylation signals and regulatory sequences that  
10 control gene expression, among others, as are known in the art. Some polypeptides require processing events, such as proteolytic cleavage, covalent modification, etc., in order to become fully active. Accordingly, functionally active nucleic acids may encode the mature or the pre-processed aconitase polypeptide, or an intermediate form. An aconitase polynucleotide can also include heterologous coding sequences, for example,  
15 sequences that encode a marker included to facilitate the purification of the fused polypeptide, or a transformation marker.

In another embodiment, a functionally active aconitase nucleic acid is capable of being used in the generation of loss-of-function aconitase phenotypes, for instance, via antisense suppression, co-suppression, etc.

20 In one preferred embodiment, an aconitase nucleic acid used in the methods of this invention comprises a nucleic acid sequence that encodes or is complementary to a sequence that encodes an aconitase polypeptide having at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95% or more sequence identity to the polypeptide sequence presented in SEQ ID NO:2.

25 In another embodiment an aconitase polypeptide of the invention comprises a polypeptide sequence with at least 50% or 60% identity to the aconitase polypeptide sequence of SEQ ID NO:2, and may have at least 70%, 80%, 85%, 90% or 95% or more sequence identity to the aconitase polypeptide sequence of SEQ ID NO:2. In another embodiment, an aconitase polypeptide comprises a polypeptide sequence with at least  
30 50%, 60%, 70%, 80%, 85%, 90% or 95% or more sequence identity to a functionally active fragment of the polypeptide presented in SEQ ID NO:2. In yet another embodiment, an aconitase polypeptide comprises a polypeptide sequence with at least 50%, 60 %, 70%, 80%, or 90% identity to the polypeptide sequence of SEQ ID NO:2 over its entire length.

In another aspect, an aconitase polynucleotide sequence is at least 50% to 60% identical over its entire length to the aconitase nucleic acid sequence presented as SEQ ID NO:1, or nucleic acid sequences that are complementary to such an aconitase sequence, and may comprise at least 70%, 80%, 85%, 90% or 95% or more sequence identity to the aconitase sequence presented as SEQ ID NO:1 or a functionally active fragment thereof, or complementary sequences.

As used herein, "percent (%) sequence identity" with respect to a specified subject sequence, or a specified portion thereof, is defined as the percentage of nucleotides or amino acids in the candidate derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul *et al.*, J. Mol. Biol. (1997) 215:403-410; website at [blast.wustl.edu/blast/README.html](http://blast.wustl.edu/blast/README.html)) with search parameters set to default values. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A "% identity value" is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation. A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the protein is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that selectively hybridize to the nucleic acid sequence of SEQ ID NO:1. The stringency of hybridization can be controlled by temperature, ionic strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing.

Conditions routinely used are well known (see, *e.g.*, Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor (1989)). In some embodiments, a nucleic acid molecule of the invention is capable of hybridizing to a nucleic acid molecule containing the nucleotide sequence of SEQ ID NO:1 under stringent hybridization conditions that are: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X Denhardt's solution, 0.05% sodium pyrophosphate and 100 µg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C in a solution containing 6X SSC, 1X Denhardt's solution, 100 µg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C for 1 h in a solution containing 0.1X SSC and 0.1% SDS (sodium dodecyl sulfate). In other embodiments, moderately stringent hybridization conditions are used that are: pretreatment of filters containing nucleic acid for 6 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 µg/ml denatured salmon sperm DNA; hybridization for 18-20 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C in a solution containing 2X SSC and 0.1% SDS. Alternatively, low stringency conditions can be used that comprise: incubation for 8 hours to overnight at 37° C in a solution comprising 20% formamide, 5 x SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1 x SSC at about 37° C for 1 hour.

As a result of the degeneracy of the genetic code, a number of polynucleotide sequences encoding an aconitase polypeptide can be produced. For example, codons may be selected to increase the rate at which expression of the polypeptide occurs in a particular host species, in accordance with the optimum codon usage dictated by the particular host organism (see, *e.g.*, Nakamura *et al*, 1999, Nucleic Acids Res 27:292). Such sequence variants may be used in the methods of this invention.

The methods of the invention may use orthologs of the *Arabidopsis* aconitase. Methods of identifying the orthologs in other plant species are known in the art. Normally, orthologs in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. In evolution, when a gene

duplication event follows speciation, a single gene in one species, such as *Arabidopsis*, may correspond to multiple genes (paralogs) in another. As used herein, the term "orthologs" encompasses paralogs. When sequence data is available for a particular plant species, orthologs are generally identified by sequence homology analysis, such as

5 BLAST analysis, usually using protein bait sequences. Sequences are assigned as a potential ortholog if the best hit sequence from the forward BLAST result retrieves the original query sequence in the reverse BLAST (Huynen MA and Bork P, Proc Natl Acad Sci (1998) 95:5849-5856; Huynen MA *et al.*, Genome Research (2000) 10:1204-1210).

Programs for multiple sequence alignment, such as CLUSTAL (Thompson JD *et al.*, 1994,

10 Nucleic Acids Res 22:4673-4680) may be used to highlight conserved regions and/or residues of orthologous proteins and to generate phylogenetic trees. In a phylogenetic tree representing multiple homologous sequences from diverse species (e.g., retrieved through BLAST analysis), orthologous sequences from two species generally appear closest on the tree with respect to all other sequences from these two species. Structural threading or

15 other analysis of protein folding (e.g., using software by ProCeryon, Biosciences, Salzburg, Austria) may also identify potential orthologs. Nucleic acid hybridization methods may also be used to find orthologous genes and are preferred when sequence data are not available. Degenerate PCR and screening of cDNA or genomic DNA libraries are common methods for finding related gene sequences and are well known in the art (see, e.g., Sambrook, *supra*; Dieffenbach C and Dveksler G (Eds.) PCR Primer: A Laboratory

20 Manual, Cold Spring Harbor Laboratory Press, NY, 1989). For instance, methods for generating a cDNA library from the plant species of interest and probing the library with partially homologous gene probes are described in Sambrook *et al., supra*. A highly conserved portion of the *Arabidopsis* aconitase coding sequence may be used as a probe.

25 Aconitase ortholog nucleic acids may hybridize to the nucleic acid of SEQ ID NO:1 under high, moderate, or low stringency conditions. After amplification or isolation of a segment of a putative ortholog, that segment may be cloned and sequenced by standard techniques and utilized as a probe to isolate a complete cDNA or genomic clone.

Alternatively, it is possible to initiate an EST project to generate a database of sequence

30 information for the plant species of interest. In another approach, antibodies that

specifically bind known aconitase polypeptides are used for ortholog isolation (see, e.g., Harlow E and Lane D, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1999, New York). Western blot analysis can determine that an aconitase ortholog (i.e., an orthologous protein) is present in a crude extract of a particular

plant species. When reactivity is observed, the sequence encoding the candidate ortholog may be isolated by screening expression libraries representing the particular plant species. Expression libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Sambrook, *et al.*, *supra*. Once the candidate  
5 ortholog(s) are identified by any of these means, candidate orthologous sequence are used as bait (the "query") for the reverse BLAST against sequences from *Arabidopsis* or other species in which aconitase nucleic acid and/or polypeptide sequences have been identified.

Aconitase nucleic acids and polypeptides may be obtained using any available method. For instance, techniques for isolating cDNA or genomic DNA sequences of  
10 interest by screening DNA libraries or by using polymerase chain reaction (PCR), as previously described, are well known in the art. Alternatively, nucleic acid sequence may be synthesized. Any known method, such as site directed mutagenesis (Kunkel TA *et al.*, *Methods Enzymol.* 204:125-39, 1991), may be used to introduce desired changes into a cloned nucleic acid.

15 In general, the methods of the invention involve incorporating the desired form of the aconitase nucleic acid into a plant expression vector for transformation of in plant cells, and the aconitase polypeptide is expressed in the host plant.

An isolated aconitase nucleic acid molecule is other than in the form or setting in which it is found in nature and is identified and separated from least one contaminant  
20 nucleic acid molecule with which it is ordinarily associated in the natural source of the aconitase nucleic acid. However, an isolated aconitase nucleic acid molecule includes aconitase nucleic acid molecules contained in cells that ordinarily express aconitase where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

### **Generation of Genetically Modified Plants with an Altered Oil Content Phenotype**

Aconitase nucleic acids and polypeptides may be used in the generation of genetically modified plants having a modified oil content phenotype. As used herein, a  
"modified oil content phenotype" may refer to modified oil content in any part of the  
30 plant; the modified oil content is often observed in seeds. In a preferred embodiment, altered expression of the aconitase gene in a plant is used to generate plants with a high oil phenotype.

The methods described herein are generally applicable to all plants. In a preferred embodiment, the invention is directed to oil-producing plants, which produce and store

triacylglycerol in specific organs, primarily in seeds. Such species include soybean (*Glycine max*), rapeseed and canola (including *Brassica napus*, *B. campestris*), sunflower (*Helianthus annuus*), cotton (*Gossypium hirsutum*), corn (*Zea mays*), cocoa (*Theobroma cacao*), safflower (*Carthamus tinctorius*), oil palm (*Elaeis guineensis*), coconut palm (*Cocos nucifera*), flax (*Linum usitatissimum*), castor (*Ricinus communis*) and peanut (*Arachis hypogaea*). The invention may also be directed to fruit- and vegetable-bearing plants, grain-producing plants, nut-producing plants, rapid cycling *Brassica* species, alfalfa (*Medicago sativa*), tobacco (*Nicotiana*), turfgrass (*Poaceae* family), other forage crops, and wild species that may be a source of unique fatty acids.

The skilled artisan will recognize that a wide variety of transformation techniques exist in the art, and new techniques are continually becoming available. Any technique that is suitable for the target host plant can be employed within the scope of the present invention. For example, the constructs can be introduced in a variety of forms including, but not limited to as a strand of DNA, in a plasmid, or in an artificial chromosome. The introduction of the constructs into the target plant cells can be accomplished by a variety of techniques, including, but not limited to *Agrobacterium*-mediated transformation, electroporation, microinjection, microprojectile bombardment calcium-phosphate-DNA co-precipitation or liposome-mediated transformation of a heterologous nucleic acid. The transformation of the plant is preferably permanent, *i.e.* by integration of the introduced expression constructs into the host plant genome, so that the introduced constructs are passed onto successive plant generations. Depending upon the intended use, a heterologous nucleic acid construct comprising an aconitase polynucleotide may encode the entire protein or a biologically active portion thereof.

In one embodiment, binary Ti-based vector systems may be used to transfer polynucleotides. Standard *Agrobacterium* binary vectors are known to those of skill in the art, and many are commercially available (e.g., pBI121 Clontech Laboratories, Palo Alto, CA).

The optimal procedure for transformation of plants with *Agrobacterium* vectors will vary with the type of plant being transformed. Exemplary methods for *Agrobacterium*-mediated transformation include transformation of explants of hypocotyl, shoot tip, stem or leaf tissue, derived from sterile seedlings and/or plantlets. Such transformed plants may be reproduced sexually, or by cell or tissue culture.

*Agrobacterium* transformation has been previously described for a large number of different types of plants and methods for such transformation may be found in the scientific

literature. Of particular relevance are methods to transform commercially important crops, such as rapeseed (De Block et al., Plant Physiol. (1989) 91:694-701), sunflower (Everett et al., Bio/Technology (1987) 5:1201), and soybean (Christou et al., Proc. Natl. Acad. Sci USA (1989) 86:7500-7504; Kline et al., Nature (1987) 327:70).

5 Expression (including transcription and translation) of aconitase may be regulated with respect to the level of expression, the tissue type(s) where expression takes place and/or developmental stage of expression. A number of heterologous regulatory sequences (e.g., promoters and enhancers) are available for controlling the expression of an aconitase nucleic acid. These include constitutive, inducible and regulatable promoters, 10 as well as promoters and enhancers that control expression in a tissue- or temporal-specific manner. Exemplary constitutive promoters include the raspberry E4 promoter (U.S. Patent Nos. 5,783,393 and 5,783,394), the 35S CaMV (Jones JD *et al.*, Transgenic Res 1:285-297 1992), the CsVMV promoter (Verdaguer B *et al.*, Plant Mol Biol 37:1055-1067, 1998) and the melon actin promoter (published PCT application WO0056863). 15 Exemplary tissue-specific promoters include the tomato E4 and E8 promoters (U.S. Pat. No. 5,859,330) and the tomato 2AII gene promoter (Van Haaren MJJ *et al.*, Plant Mol Bio 21:625-640, 1993).

In one preferred embodiment, aconitase expression is under control of regulatory sequences from genes whose expression is associated with early seed and/or embryo 20 development. Legume genes whose promoters are associated with early seed and embryo development include *V. faba legumin* (Baumlein et al., 1991, Mol Gen Genet 225:121-8; Baumlein et al., 1992, Plant J 2:233-9), *V. faba usp* (Fiedler et al., 1993, Plant Mol Biol 22:669-79), pea *convicilin* (Bown et al., 1988, Biochem J 251:717-26), pea *lectin* (dePater et al., 1993, Plant Cell 5:877-86), *P. vulgaris beta phaseolin* (Bustos et al., 1991, EMBO J 25 10:1469-79), *P. vulgaris DLEC2* and *PHS* [beta] (Bobb et al., 1997, Nucleic Acids Res 25:641-7), and soybean *beta-Conglycinin*, 7S storage protein (Chamberland et al., 1992, Plant Mol Biol 19:937-49). Cereal genes whose promoters are associated with early seed and embryo development include rice *glutelin* ("GluA-3," Yoshihara and Takaiwa, 1996, Plant Cell Physiol 37:107-11; "GluB-1," Takaiwa et al., 1996, Plant Mol Biol 30:1207-21; 30 Washida et al., 1999, Plant Mol Biol 40:1-12; "Gt3," Leisy et al., 1990, Plant Mol Biol 14:41-50), rice *prolamin* (Zhou & Fan, 1993, Transgenic Res 2:141-6), wheat *prolamin* (Hammond-Kosack et al., 1993, EMBO J 12:545-54), maize *zein* (Z4, Matzke et al., 1990, Plant Mol Biol 14:323-32), and barley *B-hordeins* (Entwistle et al., 1991, Plant Mol Biol 17:1217-31). Other genes whose promoters are associated with early seed and embryo

development include oil palm GLO7A (7S globulin, Morcillo et al., 2001, *Physiol Plant* 112:233-243), *Brassica napus napin*, 2S storage protein, and napA gene (Josefsson et al., 1987, *J Biol Chem* 262:12196-201; Stalberg et al., 1993, *Plant Mol Biol* 1993 23:671-83; Ellerstrom et al., 1996, *Plant Mol Biol* 32:1019-27), *Brassica napus oleosin* (Keddie et al., 1994, *Plant Mol Biol* 24:327-40), *Arabidopsis oleosin* (Plant et al., 1994, *Plant Mol Biol* 25:193-205), *Arabidopsis* FAE1 (Rossak et al., 2001, *Plant Mol Biol* 46:717-25), *Canavalia gladiata* conA (Yamamoto et al., 1995, *Plant Mol Biol* 27:729-41), and *Catharanthus roseus* strictosidine synthase (Str, Ouwerkerk and Memelink, 1999, *Mol Gen Genet* 261:635-43). In another preferred embodiment, regulatory sequences from genes expressed during oil biosynthesis are used (see, e.g., US Pat No: 5,952, 544). Alternative promoters are from plant storage protein genes (Bevan et al, 1993, *Philos Trans R Soc Lond B Biol Sci* 342:209-15).

In yet another aspect, in some cases it may be desirable to inhibit the expression of endogenous aconitase in a host cell. Exemplary methods for practicing this aspect of the invention include, but are not limited to antisense suppression (Smith, *et al.*, *Nature* 334:724-726, 1988; van der Krol et al., *Biotechniques* (1988) 6:958-976); co-suppression (Napoli, *et al.*, *Plant Cell* 2:279-289, 1990); ribozymes (PCT Publication WO 97/10328); and combinations of sense and antisense (Waterhouse, *et al.*, *Proc. Natl. Acad. Sci. USA* 95:13959-13964, 1998). Methods for the suppression of endogenous sequences in a host cell typically employ the transcription or transcription and translation of at least a portion of the sequence to be suppressed. Such sequences may be homologous to coding as well as non-coding regions of the endogenous sequence. Antisense inhibition may use the entire cDNA sequence (Sheehy et al., *Proc. Natl. Acad. Sci. USA* (1988) 85:8805-8809), a partial cDNA sequence including fragments of 5' coding sequence, (Cannon et al., *Plant Molec. Biol.* (1990) 15:39-47), or 3' non-coding sequences (Ch'ng et al., *Proc. Natl. Acad. Sci. USA* (1989) 86:10006-10010). Cosuppression techniques may use the entire cDNA sequence (Napoli et al., *supra*; van der Krol et al., *The Plant Cell* (1990) 2:291-299) or a partial cDNA sequence (Smith et al., *Mol. Gen. Genetics* (1990) 224:477-481).

Standard molecular and genetic tests may be performed to further analyze the association between a gene and an observed phenotype. Exemplary techniques are described below.



### 1. DNA/RNA analysis

The stage- and tissue-specific gene expression patterns in mutant versus wild-type lines may be determined, for instance, by in situ hybridization. Analysis of the methylation status of the gene, especially flanking regulatory regions, may be performed. Other suitable techniques include overexpression, ectopic expression, expression in other plant species and gene knock-out (reverse genetics, targeted knock-out, viral induced gene silencing [VIGS, see Baulcombe D, *Arch Virol Suppl* 15:189-201, 1999]).

In a preferred application expression profiling, generally by microarray analysis, is used to simultaneously measure differences or induced changes in the expression of many different genes. Techniques for microarray analysis are well known in the art (Schena M *et al.*, *Science* (1995) 270:467-470; Baldwin D *et al.*, *Cur Opin Plant Biol.* 2(2):96-103, 1999; Dangond F, *Physiol Genomics* (2000) 2:53-58; van Hal NL *et al.*, *J Biotechnol* (2000) 78:271-280; Richmond T and Somerville S, *Curr Opin Plant Biol* (2000) 3:108-116). Expression profiling of individual tagged lines may be performed. Such analysis can identify other genes that are coordinately regulated as a consequence of the overexpression of the gene of interest, which may help to place an unknown gene in a particular pathway.

### 2. Gene Product Analysis

Analysis of gene products may include recombinant protein expression, antisera production, immunolocalization, biochemical assays for catalytic or other activity, analysis of phosphorylation status, and analysis of interaction with other proteins via yeast two-hybrid assays.

### 3. Pathway Analysis

Pathway analysis may include placing a gene or gene product within a particular biochemical, metabolic or signaling pathway based on its mis-expression phenotype or by sequence homology with related genes. Alternatively, analysis may comprise genetic crosses with wild-type lines and other mutant lines (creating double mutants) to order the gene in a pathway, or determining the effect of a mutation on expression of downstream "reporter" genes in a pathway.

## **Generation of Mutated Plants with an Altered Oil Content Phenotype**

The invention further provides a method of identifying plants that have mutations in endogenous aconitase that confer altered oil content, and generating altered oil content progeny of these plants that are not genetically modified. In one method, called

“TILLING” (for targeting induced local lesions in genomes), mutations are induced in the seed of a plant of interest, for example, using EMS treatment. The resulting plants are grown and self-fertilized, and the progeny are used to prepare DNA samples. Aconitase-specific PCR is used to identify whether a mutated plant has an aconitase mutation. Plants having aconitase mutations may then be tested for altered oil content, or alternatively, plants may be tested for altered oil content, and then aconitase-specific PCR is used to determine whether a plant having altered oil content has a mutated aconitase gene. TILLING can identify mutations that may alter the expression of specific genes or the activity of proteins encoded by these genes (see Colbert et al (2001) Plant Physiol 126:480-484; McCallum et al (2000) Nature Biotechnology 18:455-457).

In another method, a candidate gene/Quantitative Trait Locus (QTLs) approach can be used in a marker-assisted breeding program to identify alleles of or mutations in the aconitase gene or orthologs of aconitase that may confer altered oil content (see Bert et al., Theor Appl Genet. 2003 Jun;107(1):181-9; and Lionneton et al, Genome. 2002 Dec;45(6):1203-15). Thus, in a further aspect of the invention, an aconitase nucleic acid is used to identify whether a plant having altered oil content has a mutation in endogenous aconitase or has a particular allele that causes altered oil content.

All publications cited herein are expressly incorporated herein by reference for the purpose of describing and disclosing compositions and methodologies that might be used in connection with the invention. All cited patents, patent applications, and sequence information in referenced websites and public databases are also incorporated by reference.

## EXAMPLES

### EXAMPLE 1

#### Generation of transgenic plants overexpressing aconitase

A subset of the genes encoding enzymes of the glyoxylate pathway were over-expressed in wild-type Arabidopsis and seed from the T1 transgenic plants were tested for a high oil phenotype. To over-express the genes, genomic DNA was PCR amplified with primers specific for each gene. The PCR product was cloned behind the CsVMV promoter in a T-DNA vector containing the NPTII gene (which serves as a selectable marker) and transformed into wild-type Arabidopsis plants. Transgenic plants were selected by germinating seeds on agar medium containing kanamycin. Kanamycin resistant seedlings were transplanted to soil and grown to maturity. Non-transformed

wild-type Col-0 plants were used as a control. Seed was germinated on agar medium (lacking kanamycin) and seedlings were transplanted to soil and grown to maturity. Oil content in seed harvested from both the transgenic and control plants was measured using Near Infrared Spectroscopy (NIR). NIR infrared spectra were captured using a Bruker 22 N/F. Bruker Software was used to estimate total seed oil content using data from NIR analysis and reference methods according to the manufacturers instructions.

The results showed that over-expression of the Arabidopsis aconitase gene (At4g35830) confers a high seed oil phenotype. Seed from 7 transgenic lines over-expressing At4g35830 (encoding aconitase) had more oil than all of the control plants tested. The values ranged from 42.2% to 41.8% oil for the transgenic plants while the values from the control seed ranged between 41.7% and 38.1%. The average oil content in the control seed was 40.3%. Thus, over-expression of aconitase can confer as much as a 5% increase in seed oil.

## EXAMPLE 2

### Analysis of Arabidopsis Aconitase

Sequence analyses were performed with BLAST (Altschul *et al.*, 1997, J. Mol. Biol. 215:403-410), PFAM (Bateman *et al.*, 1999, Nucleic Acids Res 27:260-262), PSORT (Nakai K, and Horton P, 1999, Trends Biochem Sci 24:34-6), and/or CLUSTAL (Thompson JD *et al.*, 1994, Nucleic Acids Res 22:4673-4680). Numerous orthologs were identified having high sequence identity, and thus are expected to also confer a high oil phenotype when overexpressed in a plant. Orthologous aconitases from various plant species include: GI#599625, GI#25291984, GI#18416900, GI#15215804, and GI#7437041 (*Arabidopsis thaliana*); GI#34851120 (*Prunus avium*); GI#11066033 (*Nicotiana tabacum*); GI#1351856 (*Cucurbita cv.*); GI#30407706 (*Lycopersicon pennellii*); GI#3309243 (*Citrus limon*); and GI#2492636 (*Cucumis melo*).

## IT IS CLAIMED:

1. A transgenic plant comprising a plant transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that encodes an aconitase polypeptide comprising the amino acid sequence of SEQ ID NO:2, or an ortholog thereof,  
5       whereby the transgenic plant has a high oil phenotype relative to control plants.
2. The transgenic plant of Claim 1, which is selected from the group consisting of rapeseed, soy, corn, sunflower, cotton, cocoa, safflower, oil palm, coconut palm, flax,  
10       castor and peanut.
3. A plant part obtained from the plant according to Claim 1.
4. The plant part of Claim 3, which is a seed.
- 15       5. A method of producing oil comprising growing the transgenic plant of Claim 1 and recovering oil from said plant.
6. A method of producing a high oil phenotype in a plant, said method comprising:  
20       a) introducing into progenitor cells of the plant a plant transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that encodes an aconitase polypeptide comprising the amino acid sequence of SEQ ID NO:2, or an ortholog thereof, and  
       b) growing the transformed progenitor cells to produce a transgenic plant, wherein  
25       said polynucleotide sequence is expressed, and said transgenic plant exhibits an altered oil content phenotype relative to control plants.
7. A plant obtained by a method of Claim 6.
- 30       8. The plant of Claim 7, which is selected from the group consisting of rapeseed, soy, corn, sunflower, cotton, cocoa, safflower, oil palm, coconut palm, flax, castor and peanut.
9. A method of generating a plant having a high oil phenotype comprising identifying a plant that has an allele in its aconitase gene that results in increased oil content compared

to plants lacking the allele and generating progeny of said identified plant, wherein the generated progeny inherit the allele and have the high oil phenotype.

10. The method of Claim 9 that employs candidate gene/QTL methodology.

5

11. The method of Claim 9 that employs TILLING methodology.

## SEQUENCE LISTING

&lt;110&gt; Agrinomics LLC

&lt;120&gt; GENERATION OF PLANTS WITH ALTERED OIL CONTENT

&lt;130&gt; AG03-083C-PC

&lt;150&gt; 60/434,601

&lt;151&gt; 2002-12-18

&lt;160&gt; 2

&lt;170&gt; PatentIn version 3.2

&lt;210&gt; 1

&lt;211&gt; 3214

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 1

```

gaatcggctc aggctcgtgt agcctcacca actccgctta taaactttct ctcttctgaa      60
aacagatctc tctactttgc ttcttcactc gacttgtatc tatcatccat ggcttccgag      120
aatcctttcc gaagcatatt gaaggcggtta gagaagcctg atggtggtga attcggtaac      180
tactacagct tacctgcttt gaacgatccc aggatcgata aactacctta ttccattagg      240
atacttcttg aatcggccat acgtaactgt gatgagttcc aagttaagag caaagatggt      300
gagaagattc ttgattggga gaatacttct cccaagcagg ttgagattcc gttcaagcct      360
gctcgggttc ttcttcagga ctttactggt gttcctgctg ttgttgatct tgcttgcatg      420
agagatgcca tgaataatct cgggtggtgat tctaataaaa ttaatccgct ggtccctgta      480
gatcttgcca ttgatcactc cgttcagggtg gatgtggcga gatcagagaa cgcagtgcag      540
gcaaacatgg agcttgagtt ccagcgtaac aaggaaagat ttgcttttct taagtgggga      600
tccaacgcct ttcacaacat gcttgctgta cctcctggat ctggaatagt tcatcaagtc      660
aacctagaat accttgccag agttgttttc aacacaaatg gacttcttta ccagacagt      720
gttggtggca cagactctca caccactatg attgatggac tgggtggtgc tggatgggga      780
gttggcggta tagaagcgga agctaccatg cttggtcagc caatgagcat ggtcctaccc      840
ggtgttgtgg gtttcaagct aacgggaaag ttaagagatg gaatgacagc tactgatttg      900
gtcttaacag tgactcagat gttgaggaaa catggagtag ttggaaagtt tgttgaattc      960
cacggggaag ggatgagaga attgtcttta gctgaccgtg ctacaattgc caatatgtct     1020
cctgagtacg gtgcgaccat gggattcttc ccagtcgata atgtcacttt gcagtatcta     1080
aggttgacag gcaggagcga tgacactgtc tccatgatag aggcgtatct acgagcaaac     1140
aagatgtttg tggattacag tgagccggag agtaagacag tttattcctc atgtctggaa     1200
ttgaatctcg aggatgtgga accttggtgt tctggtccca agaggcctca tgatcgtggt     1260

```

cctttgaagg aaatgaaagc ggactggcat tcttgcttgg acaatagagt aggattcaag 1320  
ggtttcgctg tacctaaaga agcacagagt aaggctgtag agttcaattt taacgggacc 1380  
acagcacagc ttagacatgg agatgttggt atagcagcaa tcaccagttg cacaataact 1440  
tcaaacccta gtgtaatgct tggcgctgcc ttagttgcaa aaaaggcctg cgacctagga 1500  
ctggagggtta agccatggat caaaactagt cttgctccag gctctggagt tgtaacaaag 1560  
tacttggcaa agagtggctt gcagaagtac ttgaatcagc tcggcttcag tatcgttggt 1620  
tatgggtgca ccacatgcat tggaaactcg ggggatatcc atgaagctgt ggcttcagca 1680  
atagttgata atgacttggt ggcacccgt gtgttgctct ggaacagaaa ttttgaggga 1740  
cgtgttcacc cgtaacaag agctaactat ctagcttccc caccgcttgt ttagcctat 1800  
gctctggctg gaactgttga cattgatttt gagacacagc ccattggaac tgggaaagat 1860  
ggaaaacaga tatttttcag ggacatttg cctctaaca aagaagttgc tgaggttgtt 1920  
caatctagtg tccttcctga tatgttcaa gctacatat aagcaatcac caaaggaaat 1980  
tccatgtgga atcagttatc tgtggcgta ggtactctct atgagtggga cccgaaatca 2040  
acttacattc acgagccgcc ttatttcaag ggcatgacca tgtctccacc cgggccacat 2100  
ggtgtgaaag acgcatactg ttactcaat tttggagaca gtattaccac tgatcacatc 2160  
tcaccagctg gtagcatcca caaggacagt cctgcggcta agtacttgat ggaacgaggt 2220  
gtggatagaa gagacttcaa ctcatcggg agtcgccgtg gtaatgatga gattatggcg 2280  
agaggcactt ttgcaaatat ccgtattgtc aacaaacact tgaaaggaga agttgggtccc 2340  
aaaacagttc acattcccac tggagagaag ctttctgttt tcgatgctgc catgaaatat 2400  
aggaacgagg gacgcgacac aatcattttg gctgggtgctg aatacggtag tggaagttct 2460  
cgtgattggg ctgccaaggg tccaatgctt ctgggtgtga aagctgtgat ttcaaagagc 2520  
ttcgagcgaa ttcaccgaag caatttggtg ggaatgggaa tcataccttt gtgcttcaag 2580  
gcgggagaag atgctgagac cttggccta acgggtcagg agctttacac cattgagctc 2640  
ccaacaatg ttagtgagat caaaccagga caagatgtaa cagtcgtcac aaacaatggc 2700  
aaatctttca catgtacact ccgatttgac acagaggtgg agttggctta tttcgatcac 2760  
ggagggattt tgcaatacgt tatcaggaac ttgatcaaac aataaatctg gtaacaccag 2820  
agacttgagt tatatatccc aaggtttggt gcaataaaaa tgtttgcagg gaggaggacg 2880  
agaatgactt taatttaaac ttttgctttt gttcttcttc gtctcttgct tcgtgctagt 2940  
caattggaaa cattatcact gtcgagtcat ttttttctt tcaaataaac atgcgagctc 3000  
tttttttggt gttgttggtc ataagcattg tcaatgctgc attgatagaa gtatactcta 3060  
ccattagaaa ataaacacaa atacacaaag taatctaaag agctagagga tgaaaattat 3120

cttgtgaagg ttgtacaaaa acattaaaaa aaatctgaga tgatccaaag gattgattat 3180  
 cacattcgaa gtgatgggtca gagattacct cttg 3214

<210> 2  
 <211> 898  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 2

Met Ala Ser Glu Asn Pro Phe Arg Ser Ile Leu Lys Ala Leu Glu Lys  
 1 5 10 15

Pro Asp Gly Gly Glu Phe Gly Asn Tyr Tyr Ser Leu Pro Ala Leu Asn  
 20 25 30

Asp Pro Arg Ile Asp Lys Leu Pro Tyr Ser Ile Arg Ile Leu Leu Glu  
 35 40 45

Ser Ala Ile Arg Asn Cys Asp Glu Phe Gln Val Lys Ser Lys Asp Val  
 50 55 60

Glu Lys Ile Leu Asp Trp Glu Asn Thr Ser Pro Lys Gln Val Glu Ile  
 65 70 75 80

Pro Phe Lys Pro Ala Arg Val Leu Leu Gln Asp Phe Thr Gly Val Pro  
 85 90 95

Ala Val Val Asp Leu Ala Cys Met Arg Asp Ala Met Asn Asn Leu Gly  
 100 105 110

Gly Asp Ser Asn Lys Ile Asn Pro Leu Val Pro Val Asp Leu Val Ile  
 115 120 125

Asp His Ser Val Gln Val Asp Val Ala Arg Ser Glu Asn Ala Val Gln  
 130 135 140

Ala Asn Met Glu Leu Glu Phe Gln Arg Asn Lys Glu Arg Phe Ala Phe  
 145 150 155 160

Leu Lys Trp Gly Ser Asn Ala Phe His Asn Met Leu Val Val Pro Pro  
 165 170 175

Gly Ser Gly Ile Val His Gln Val Asn Leu Glu Tyr Leu Ala Arg Val  
 180 185 190

Val Phe Asn Thr Asn Gly Leu Leu Tyr Pro Asp Ser Val Val Gly Thr



195	200	205
Asp Ser His Thr Thr Met Ile Asp Gly Leu Gly Val Ala Gly Trp Gly		
210	215	220
Val Gly Gly Ile Glu Ala Glu Ala Thr Met Leu Gly Gln Pro Met Ser		
225	230	235 240
Met Val Leu Pro Gly Val Val Gly Phe Lys Leu Thr Gly Lys Leu Arg		
	245	250 255
Asp Gly Met Thr Ala Thr Asp Leu Val Leu Thr Val Thr Gln Met Leu		
	260	265 270
Arg Lys His Gly Val Val Gly Lys Phe Val Glu Phe His Gly Glu Gly		
	275	280 285
Met Arg Glu Leu Ser Leu Ala Asp Arg Ala Thr Ile Ala Asn Met Ser		
	290	295 300
Pro Glu Tyr Gly Ala Thr Met Gly Phe Phe Pro Val Asp His Val Thr		
305	310	315 320
Leu Gln Tyr Leu Arg Leu Thr Gly Arg Ser Asp Asp Thr Val Ser Met		
	325	330 335
Ile Glu Ala Tyr Leu Arg Ala Asn Lys Met Phe Val Asp Tyr Ser Glu		
	340	345 350
Pro Glu Ser Lys Thr Val Tyr Ser Ser Cys Leu Glu Leu Asn Leu Glu		
	355	360 365
Asp Val Glu Pro Cys Val Ser Gly Pro Lys Arg Pro His Asp Arg Val		
	370	375 380
Pro Leu Lys Glu Met Lys Ala Asp Trp His Ser Cys Leu Asp Asn Arg		
385	390	395 400
Val Gly Phe Lys Gly Phe Ala Val Pro Lys Glu Ala Gln Ser Lys Ala		
	405	410 415
Val Glu Phe Asn Phe Asn Gly Thr Thr Ala Gln Leu Arg His Gly Asp		
	420	425 430
Val Val Ile Ala Ala Ile Thr Ser Cys Thr Asn Thr Ser Asn Pro Ser		
	435	440 445

Val Met Leu Gly Ala Ala Leu Val Ala Lys Lys Ala Cys Asp Leu Gly  
 450 455 460

Leu Glu Val Lys Pro Trp Ile Lys Thr Ser Leu Ala Pro Gly Ser Gly  
 465 470 475 480

Val Val Thr Lys Tyr Leu Ala Lys Ser Gly Leu Gln Lys Tyr Leu Asn  
 485 490 495

Gln Leu Gly Phe Ser Ile Val Gly Tyr Gly Cys Thr Thr Cys Ile Gly  
 500 505 510

Asn Ser Gly Asp Ile His Glu Ala Val Ala Ser Ala Ile Val Asp Asn  
 515 520 525

Asp Leu Val Ala Ser Ala Val Leu Ser Gly Asn Arg Asn Phe Glu Gly  
 530 535 540

Arg Val His Pro Leu Thr Arg Ala Asn Tyr Leu Ala Ser Pro Pro Leu  
 545 550 555 560

Val Val Ala Tyr Ala Leu Ala Gly Thr Val Asp Ile Asp Phe Glu Thr  
 565 570 575

Gln Pro Ile Gly Thr Gly Lys Asp Gly Lys Gln Ile Phe Phe Arg Asp  
 580 585 590

Ile Trp Pro Ser Asn Lys Glu Val Ala Glu Val Val Gln Ser Ser Val  
 595 600 605

Leu Pro Asp Met Phe Lys Ala Thr Tyr Glu Ala Ile Thr Lys Gly Asn  
 610 615 620

Ser Met Trp Asn Gln Leu Ser Val Ala Ser Gly Thr Leu Tyr Glu Trp  
 625 630 635 640

Asp Pro Lys Ser Thr Tyr Ile His Glu Pro Pro Tyr Phe Lys Gly Met  
 645 650 655

Thr Met Ser Pro Pro Gly Pro His Gly Val Lys Asp Ala Tyr Cys Leu  
 660 665 670

Leu Asn Phe Gly Asp Ser Ile Thr Thr Asp His Ile Ser Pro Ala Gly  
 675 680 685

Ser Ile His Lys Asp Ser Pro Ala Ala Lys Tyr Leu Met Glu Arg Gly

690

695

700

Val Asp Arg Arg Asp Phe Asn Ser Tyr Gly Ser Arg Arg Gly Asn Asp  
705 710 715 720

Glu Ile Met Ala Arg Gly Thr Phe Ala Asn Ile Arg Ile Val Asn Lys  
725 730 735

His Leu Lys Gly Glu Val Gly Pro Lys Thr Val His Ile Pro Thr Gly  
740 745 750

Glu Lys Leu Ser Val Phe Asp Ala Ala Met Lys Tyr Arg Asn Glu Gly  
755 760 765

Arg Asp Thr Ile Ile Leu Ala Gly Ala Glu Tyr Gly Ser Gly Ser Ser  
770 775 780

Arg Asp Trp Ala Ala Lys Gly Pro Met Leu Leu Gly Val Lys Ala Val  
785 790 795 800

Ile Ser Lys Ser Phe Glu Arg Ile His Arg Ser Asn Leu Val Gly Met  
805 810 815

Gly Ile Ile Pro Leu Cys Phe Lys Ala Gly Glu Asp Ala Glu Thr Leu  
820 825 830

Gly Leu Thr Gly Gln Glu Leu Tyr Thr Ile Glu Leu Pro Asn Asn Val  
835 840 845

Ser Glu Ile Lys Pro Gly Gln Asp Val Thr Val Val Thr Asn Asn Gly  
850 855 860

Lys Ser Phe Thr Cys Thr Leu Arg Phe Asp Thr Glu Val Glu Leu Ala  
865 870 875 880

Tyr Phe Asp His Gly Gly Ile Leu Gln Tyr Val Ile Arg Asn Leu Ile  
885 890 895

Lys Gln

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
8 July 2004 (08.07.2004)

PCT

(10) International Publication Number  
**WO 2004/056967 A3**

(51) International Patent Classification<sup>7</sup>: **A01H 5/00**,  
C12N 15/82

(21) International Application Number:  
PCT/US2003/040987

(22) International Filing Date:  
18 December 2003 (18.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/434,601 18 December 2002 (18.12.2002) US

(71) Applicant (for all designated States except US): **AGRI-  
NOMICS LLC** [US/US]; 16160 S.W. Upper Boones Ferry  
Road, Portland, OR 97224-7744 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LIGHTNER,  
Jonathan** [US/US]; 28750 S. Heisinger Lane, Mulino,  
OR 97042 (US). **COATE, Jeremy, E.** [US/US]; 3202 S.E.  
10th Avenue, Portland, OR 97202 (US). **CLENDEN-  
NEN, Stephanie, K.** [US/US]; P.O. Box 6332, Kingsport,  
TN 37663 (US). **FEDERSPIEL, Nancy, Anne** [US/US];  
1345 Altschul Avenue, Menlo Park, CA 94025 (US).

(74) Agents: **BRUNELLE, Jan, P. et al.**; **EXELIXIS, INC.**,  
170 Harbor Way, P.O. Box 511, South San Francisco, CA  
94083-0511 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD,  
GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,  
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,  
MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU,  
SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (BW, GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,  
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,  
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

- with international search report
- before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

(88) Date of publication of the international search report:  
12 August 2004

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: GENERATION OF PLANTS WITH ALTERED OIL CONTENT

(57) Abstract: The present invention is directed to plants that display an altered oil content phenotype due to altered expression of an aconitase nucleic acid. The invention is further directed to methods of generating plants with an altered oil content phenotype.

WO 2004/056967 A3

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/40987

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A01H 5/00; C12N 15/82

US CL : 800/281, 298

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/281, 298

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Sequence databases, WEST, DIALOG

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95/20046 A1 (BIOCEM) 27 July 1995 (27.05.95), see especially the abstract and Claim 15.	1-8

☐

Further documents are listed in the continuation of Box C.

☐

See patent family annex.

• Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"I" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

21 May 2004 (21.05.2004)

Date of mailing of the international search report

17 JUN 2004

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, Virginia 22313-1450

Facsimile No. (703) 305-3230

Authorized officer

Elizabeth F. McElwain

Telephone No. 703-308-0196

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/40987

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:  
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-8

Remark on Protest

☐  
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

PCT/US03/40987

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-8, drawn to a transgenic plant transformed with DNA encoding SEQ ID NO: 2, and a method of producing oil and producing a high oil phenotype in said plant.

Group II, claim(s) 9-11, drawn to a method of generating a plant with a high oil phenotype by identifying a plant having an aconitase gene that results in increased oil content, and further requires QTL or TILLING methodology.

The inventions listed as Groups I and II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I requires an isolated DNA encoding the amino acid sequence of SEQ ID NO: 2, which is not required by Group II. In addition, the methods of Group I and II require different method steps and result in different products and are of different scope with regard to alleles of aconitase, wherein Group II is not limited to a plant having in its genome DNA that encodes an amino acid sequence of SEQ ID NO: 2.